Risk Aversion or Mistaken Beliefs?[1]

Thomas J. Sargent

September 18, 2017

[1] Originally, "Three Martingales and More"

Risks and beliefs

$$
\begin{aligned}
\phi(\varepsilon) &= K \exp\left(-\frac{1}{2}\varepsilon'\varepsilon\right) \sim \mathcal{N}(0, I) \\
\hat{\phi}(\varepsilon) &= K \exp\left(-\frac{1}{2}(\varepsilon-\lambda)'(\varepsilon-\lambda)\right) \sim \mathcal{N}(-\lambda, I) \\
&= m(\varepsilon)\phi(\varepsilon) \\
m(\varepsilon) &= \exp\left(-\lambda'\varepsilon - \frac{1}{2}\lambda'\lambda\right) \geq 0 \\
Em(\varepsilon) &= 1 \\
\text{entropy} &\equiv Em(\varepsilon)\log m(\varepsilon) = \frac{1}{2}\lambda'\lambda
\end{aligned}
$$

Likelihood ratios (martingale increments)

Probability	Likelihood ratio	Describes
econometric	1	macro risk factors
risk neutral	m^{λ}_{t+1}	prices of risks
mistaken	m^{w}_{t+1}	experts' forecasts
dubious	$m_{t+1} \in \mathcal{M}$	specification doubts

$$m^b_{t+1} = \exp\left(-b_t'\varepsilon_{t+1} - \frac{b_t'b_t}{2}\right);\ b_t = 0 \text{ or } \lambda_t \text{ or } w_t \text{ or } \ldots$$

Young researchers

- Jaroslav Borovička
- Anmol Bhandari
- Timothy Christensen
- Bálint Szőke

Middle aged researchers

- Monika Piazzesi
- Jessica Wachter
- Amir Yaron
- Stanley Zin
- Lars Hansen

Rational Expectations

1970's-1980's informal justification:
RE as outcome of learning from an infinite history

- Least squares learning converges to rational expectations equilibrium
- Depends on assumption that agents know correct functional forms
- Proof technique: stochastic approximation – partition dynamics into fast (justifying a LLN) and slow (justifying an ODE)
- But in systems with long intertemporal dependence, rates of convergence are slow

Good econometricians

- Have only limited data and hunches about functional forms
- After best econometrics, they fear models are incorrect
- Oliver Cromwell's rule: "think it possible that you might be mistaken"

Agent like good econometrician

- Has parametric model estimated from limited data
- Acknowledges that other specifications fit nearly as well
    - Other parameter values
    - Other functional forms
    - Other nonlinearities and history dependencies

Econometrician's and agent's shared model

$$
\begin{aligned}
x_{t+1} &= Ax_t + C\varepsilon_{t+1} \\
y_{t+1} &= Dx_t + G\varepsilon_{t+1} \\
\varepsilon_{t+1} &\sim \mathcal{N}(0, I) \\
r_t &= \bar{r}x_t \\
d_t &= \bar{d}x_t
\end{aligned}
$$

y_{t+1}: utility-relevant variables
r_t: risk-free one-period interest rate
d_t: payout process from an asset

Want

Price at t of a claim to random payout stream $\{d_{t+j}\}_{j=1}^{\infty}$

Rational expectations with risk neutral representative investor

Stock prices (Shiller):
Stock price p_t:

$$p_t = \exp(-r_t) E_t(p_{t+1} + d_{t+1})$$

Expectations theory of term structure of interest rates (Hicks-Shiller):
time t price $p_t(n)$ of a zero coupon n risk-free claim to one dollar at time $t+n$

$$\begin{aligned} p_t(1) &= \exp(-r_t) \\ p_t(n+1) &= \exp(-r_t) E_t p_{t+1}(n) \\ p_t(n) &= \exp(B_n x_t) \end{aligned}$$

Rational expectations with risk neutral representative investor

Works

- “Pretty well” for conditional means
- Less well for conditional variances (Shiller “volatility puzzles”)

Wouldn't it be nice . . .

. . . if we could make the theory apply even if investors

- Are risk averse
- Don't have rational expectations

but continue to use the same formulas

"Ask and it shall be given" (if you don't ask too much)

Likelihood ratio

Let

$$
\begin{aligned}
m_{t+1} &= \exp\left(-\lambda_t \varepsilon_{t+1} - \frac{1}{2}\lambda_t' \lambda_t\right) \\
\lambda_t &= \lambda x_t \\
E_t m_{t+1} &= 1, \quad m_{t+1} \geq 0
\end{aligned}
$$

- m_{t+1} is a likelihood ratio that distorts conditional distribution of ε_{t+1}.
- Multiplication of $\mathcal{N}(0, I)$ by m_{t+1} shifts density of ε_{t+1} to $\mathcal{N}(-\lambda x_t, I)$.
- Covariances of returns with m_{t+1} affect mean returns

Likelihood ratio, II

- Likelihood ratio expresses risk aversion
- λ_t is price representative agent charges for bearing exposure to ε_{t+1}
- Expected return of an asset depends on how its payout covaries with ε_{t+1}

Modern (post Shiller) asset pricing

Stock price (Lucas-Hansen):

$$p_t = \exp(-r_t) E_t(m_{t+1}(p_{t+1} + d_{t+1}))$$

Term structure (Backus-Zin):

$$\begin{aligned} p_t(1) &= \exp(-r_t) \\ p_t(n+1) &= \exp(-r_t) E_t(m_{t+1} p_{t+1}(n)) \\ p_t(n) &= \exp(B_n x_t) \end{aligned}$$

"Ask and it shall be given" works (this time)

"Risk-neutral" dynamics

$$
\begin{aligned}
x_{t+1} &= (A - C\lambda)x_t + C\tilde{\varepsilon}_{t+1} \\
\tilde{\varepsilon}_{t+1} &\sim \mathcal{N}(0, I) \\
\lambda_t &= \lambda x_t
\end{aligned}
$$

Risk-neutral dynamics

- Risk neutral dynamics assert that shock distribution ε_{t+1} has conditional mean $-\lambda_t$ instead of 0.
- Dependence of λ_t on x_t modifies dynamics asserted by econometrician's model

Expectation under twisted distribution

Mathematical expectation of y_{t+1} under probability distribution twisted by likelihood ratio m_{t+1} is

$$\tilde{E}_t y_{t+1} = E_t m_{t+1} y_{t+1}$$

Risk-neutral pricing represented

With respect to risk neutral dynamics, modern (Backus-Zin) term structure theory is

$$\begin{aligned} p_t(1) &= \exp(-r_t) \\ p_t(n+1) &= \exp(-r_t)\tilde{E}_t(p_{t+1}(n)) \\ p_t(n) &= \exp(B_n x_t) \end{aligned}$$

where $\tilde{E}_t$ is an expectation with respect to the risk-neutral measure.

- Same formulas as (Shiller) rational expectations asset pricing theory, but ...
- Take mathematical expectations with respect to a different probability measure

Another likelihood ratio m_{t+1}

Mistaken beliefs (according to the econometrician's model):

- Identical asset pricing formulas
- Identical econometric fits

Insight of Hansen, Sargent, Tallarini (1999) and Piazzesi, Salomao, Schneider (2015)

Identification

λ_t can be interpreted as either

- Risk price vector expressing "representative" agent's risk aversion, or
- Representative agent's belief distortion relative to econometrician's model

To distinguish, need more information (Piazzesi, Salomao, Schneider, PSS) or more theory (Hansen and Szőke) or both (Szőke)

More information: experts' forecasts

Piazzesi, Salomao, Schneider (2015)

- Representative agent's risk aversion leads him to price risks ε_{t+1} with prices $\lambda_t^* = \lambda^* x_t$
- Representative agent has twisted beliefs $(A^*, C) = (A - Cw^*, C)$ relative to econometrician's model (A, C)
- Professional forecasters use twisted beliefs (A^*, C) to answer survey questions about forecasts

More information: experts' forecasts

Piazzesi, Salomao, Schneider (2015)

- Use data on $\{x_t\}_{t=0}^{T}$ to estimate econometrician's model A, C
- Project experts' forecasts $\{\hat{x}_{t+1}\}$ to get $\hat{x}_{t+1} = A^* x_t$ and interpret A^* as belief distortion
- Back out mean distortion $w^* x_t = -C^{-1}(A^* - A)x_t$ to density of ε_{t+1}
- Reinterpret λ estimated by rational expectations econometrician as $\lambda^* + w^*$, where $\lambda_t^* = \lambda^* x_t$ is the (smaller) price of risk vector actually charged by a distorted beliefs representative agent

PSS approach

An econometrician who mistakenly imposes rational expectations estimates risk prices λ_t that consist of the sum of two parts:

- Smaller risk prices λ_t^* actually charged by the twisted beliefs representative agent
- Conditional mean distortions w_t^* of the risks ε_{t+1} that the twisted beliefs representative agent's model displays relative to the econometrician's

PSS empirical specification

Key variables in state space system

- Level and slope of the yield curve
  - short rate
  - spread between 5 year and short rate
- Inflation
- Conditional expectations of these variables

PSS Successes

$w^* \neq 0$

- Experts' model differs systematically from econometrician's
- Experts perceive level and slope of yield curve to be more persistent than the econometrician estimates
- Subjective risk prices $\lambda^* x_t$ vary less than λx_t estimated by rational expectations econometrician

Why are beliefs distorted?

- PSS offer no explanation
- Mistakes? Ignorance of good econometrics? Or ???
- How distorted are they?

A theory of belief distortions

A dubious investor

- Shares the econometrician's model but doubts it.
- Evaluates streams of payouts under a (vast) set of alternative specifications near his model (which equals the econometrician's)
- Constructs lower bound on set of values traced out by a set of models

Valuation under econometrician's model

Log consumption process

$$\begin{aligned} c_{t+1} - c_t &= Dx_t + G\varepsilon_{t+1} \\ x_{t+1} &= Ax_t + C\varepsilon_{t+1} \end{aligned}$$

Value function

$$V(x_0, c_0) := E\left[\sum_{t=0}^{\infty} \beta^t c_t \mid x_0, c_0\right] = c_0 + \beta E\left[V(x_1, c_1) \mid x_0, c_0\right]$$

Lars Hansen's dubious agent

- Shares econometrician's model A, C, D, G
- Expresses doubts by using (a continuum of) likelihood ratios to form discounted entropy ball of size η around econometrician's model.
- Wants valuation that is good for every model in the entropy ball.
- Constructs lower bound on values and worst-case model that attains it

Why discounted entropy?

It includes models that undiscounted entropy excludes

- Undiscounted entropy over infinite sequences excludes many models that are very difficult to distinguish from econometrician's model with limited data
- Undiscounted entropy includes only models that share laws of large numbers

Hansen agent's sequence problem, I

$$
\begin{aligned}
J(x_0, c_0 \mid \eta) := \min_{\{m_{t+1}\}_{t=0}^{\infty}} \; & E\left[\sum_{t=0}^{\infty} \beta^t M_t c_t \mid x_0, c_0\right] \\
\text{s.t.} \quad & c_{t+1} - c_t = D x_t + G \varepsilon_{t+1} \qquad \varepsilon_{t+1} \sim \mathcal{N}(0, I) \\
& x_{t+1} = A x_t + C \varepsilon_{t+1} \\
& E\left[\sum_{t=0}^{\infty} \beta^t M_t E\left[m_{t+1} \log m_{t+1} \mid x_t, c_t\right] \mid x_0, c_0\right] \leq \eta \\
& M_{t+1} = M_t m_{t+1}, \quad E[m_{t+1} \mid x_t, c_t] = 1, \quad M_0 = 1
\end{aligned}
$$

Likelihood ratio process $\{M_t\}_{t=0}^{\infty}$ is a multiplicative martingale

Entropy

Likelihood ratio

$$m_{t+1} := \exp\left(-\frac{w_t' w_t}{2} - w_t' \varepsilon_{t+1}\right)$$

implies

$$E\left[m_{t+1} \log m_{t+1} \mid x_t, c_t\right] = \frac{1}{2} w_t' w_t$$

Simplifies dubious agent's Bellman equation

Hansen agent's sequence problem, II

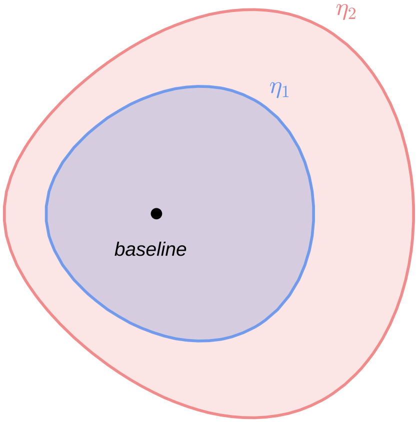
$$J(x_0, c_0 \mid \eta) := \min_{\{w_t\}_{t \geq 1}} E^w \left[\sum_{t=0}^{\infty} \beta^t c_t \mid x_0, c_0 \right]$$

$$\text{s.t.} \quad c_{t+1} - c_t = Dx_t + G(\tilde{\varepsilon}_{t+1} - w_t), \qquad \tilde{\varepsilon}_{t+1} \sim \mathcal{N}(0, I)$$

$$x_{t+1} = Ax_t + C(\tilde{\varepsilon}_{t+1} - w_t)$$

$$\frac{1}{2} E^w \left[\sum_{t=0}^{\infty} \beta^t w_t' w_t \mid x_0, c_0 \right] \leq \eta \qquad // \ \tilde{\theta}$$

Discounted entropy ball



Szőke's dubious agent

- Shares the econometrician's model A, C, D, G
- Expresses doubts by using (a continuum of) likelihood ratios to form a discounted entropy ball around econometrician's model
- Insists that some martingales in discounted entropy ball represent particular alternative *parametric* models.
- Computes a worst-case model that attains a bound on values over this set of models.

Concern about another parametric model

Investor wants to include particular alternative model with

$$E_t \left[\bar{m}_{t+1} \log \bar{m}_{t+1} \right] = \frac{1}{2} \bar{w}_t' \bar{w}_t = \xi(x_t)$$

and discounted entropy

$$E^{\bar{w}} \left[\sum_{t=0}^{\infty} \beta^t \xi(x_t) \mid x_0, c_0 \right]$$

Replace entropy constraint with

$$\frac{1}{2} E^{w} \left[\sum_{t=0}^{\infty} \beta^t w_t' w_t \mid x_0, c_0 \right] \leq E^{w} \left[\sum_{t=0}^{\infty} \beta^t \xi(x_t) \mid x_0, c_0 \right]$$

Yaron's bazooka

Bansal and Yaron (2004) mix LRR with EZ preferences

- LRR is statistically difficult to detect and estimate; but ...
- Epstein-Zin or dubious agent really hates LRR
  - There are conditions under which EZ value function is indirect utility function of dubious agent
- That sets stage for big risk-prices

Concern about bigger "long-run risk" in Szőke model

Inspired by Bansal and Yaron (2004) LRR, an agent fears a particular

$$x_{t+1} = \bar{A} x_t + C \widetilde{\varepsilon}_{t+1}$$

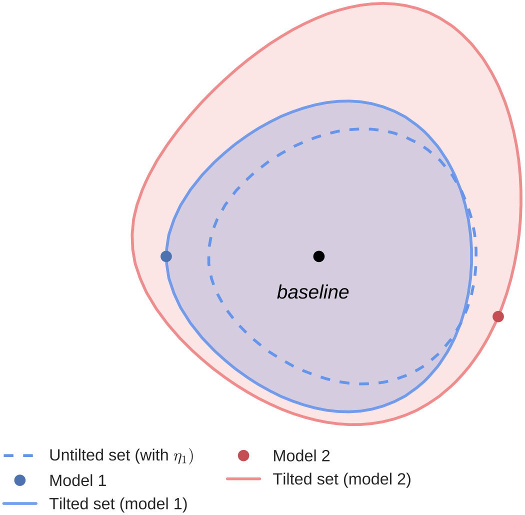
- Corresponds to $\bar{w}_t = \bar{w} x_t$ with

$$\bar{w} = -C^{-1}(\bar{A} - A)$$

- Implies quadratic ξ function:

$$\xi(x_t) := x_t' \bar{w}' \bar{w} x_t =: x_t' \Xi x_t$$

Tilted discounted entropy balls



State-dependent contributions to entropy constraint

Time t contributions to RHS of

$$\frac{1}{2} E^w \left[\sum_{t=0}^{\infty} \beta^t w_t' w_t \mid x_0, c_0 \right] \leq E^w \left[\sum_{t=0}^{\infty} \beta^t \xi(x_t) \mid x_0, c_0 \right]$$

relax the discounted entropy constraint in states x_t in which $\xi(x_t)$ is larger

This sets the stage for state-dependent mean distortions in worst-case model

That can ignite countercyclical market prices of uncertainty

Szőke agent's sequence problem

Linear quadratic problem

$$J(x_0, c_0 \mid \Xi) := \max_{\tilde{\theta} \geq 0} \min_{\{w_t\}_{t \geq 1}} E^w \left[\sum_{t=0}^{\infty} \beta^t c_t + \tilde{\theta} \frac{1}{2} \sum_{t=0}^{\infty} \beta^t \left(w_t' w_t - x_t' \Xi x_t \right) \mid x_0, c_0 \right]$$

$$\text{s.t.} \quad c_{t+1} - c_t = D x_t + G(\tilde{\varepsilon}_{t+1} - w_t), \qquad \tilde{\varepsilon}_{t+1} \sim \mathcal{N}(0, I)$$

$$x_{t+1} = A x_t + C(\tilde{\varepsilon}_{t+1} - w_t)$$

Worst-case shock mean distortion

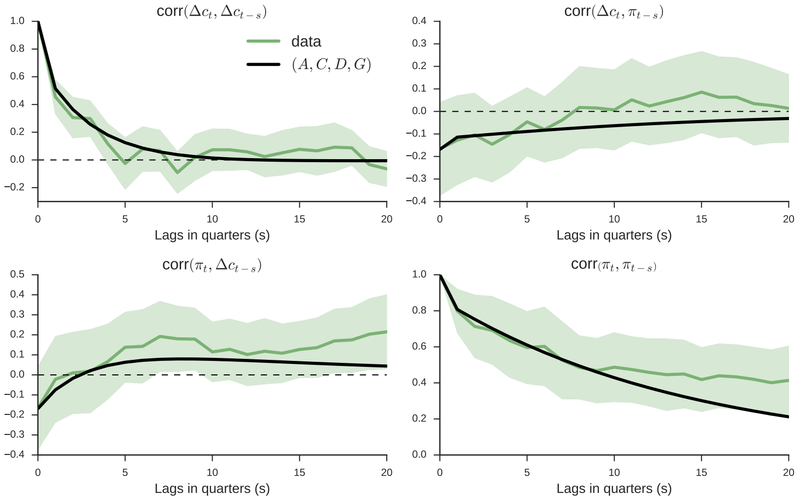
$$\widetilde{w}_t = \widetilde{w} x_t$$

Worst-case model is $(\widetilde{A}, C, \widetilde{D}, G)$

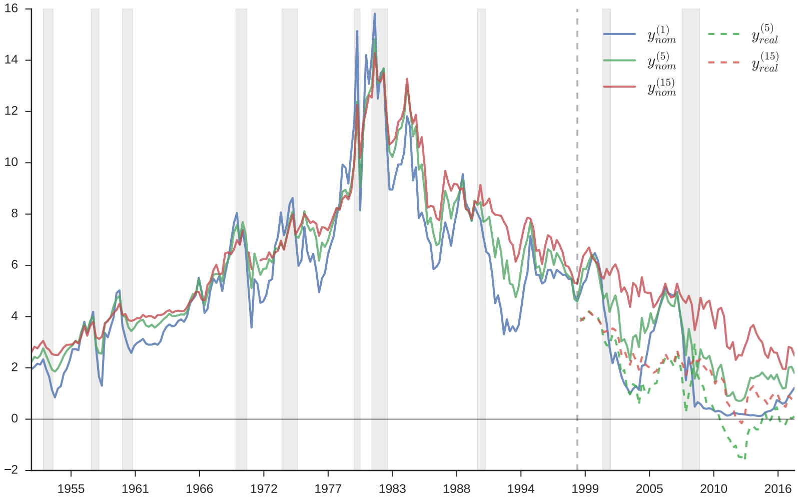
$$\widetilde{A} = A - C\widetilde{w}$$

$$\widetilde{D} = D - G\widetilde{w}$$

Econometrician’s model



US Term structure



Recognized patterns

- Nominal yield curve usually slopes upward
- Long minus short yield spread narrows before US recessions, widens after
- Consequently, slope of yield curve helps predict aggregate inputs and outputs
- Long and short yields are (almost) equally volatile ("Shiller puzzle")
- To solve "Shiller puzzle": risk prices (or something observationally equivalent) must depend on volatile state variables

Challenges and responses

	Average slope	Slopes near recessions	Volatile long yield
Lucas (1978)	no	no	no
Epstein-Zin with LRR (PS (2007), HS (2001))	maybe	yes	no
PSS (2015)	built-in	built-in	yes
Szőke (2017)	YES	yes	yes

Forces at play

- Affine risk prices with persistent consumption *growth* can nail the 2nd column
- 3rd column requires state-dependent prices of risk

Three probability twisters

- $w_t^* \sim$ Piazzesi, Salomao, Schneider's mistaken agent
- $\bar{w}_t \sim$ Szőke's especial LRR parametric worry
- $\tilde{w}_t \sim$ Szőke's worst-case model

Motivation

An appealing feature of robust control theory is that it lets us deviate from rational expectations, but still preserves a set of powerful cross-equation restrictions on decision makers' beliefs ... Consequently, estimation can proceed essentially as with rational expectations econometrics. The main difference is that now restrictions through which we interpret the data emanate from the decision maker's best response to a worst-case model instead of to the econometrician's model. Szőke (2017)

Szőke's empirical strategy, I

- Use $\{x_t, c_t\}_{t=0}^{T}$ to estimate the econometrician's A, C, D, G
- View Ξ as matrix of additional free parameters and estimate them simultaneously with risk prices $\tilde{\lambda} x_t$ in $\tilde{\lambda}_t = \tilde{\lambda} x_t$ from data $\{p_t(n+1)\}_{n=1}^{N}, t = 0, \ldots, T$ by imposing cross-equation restrictions

$$p_t(n+1) = \exp(-r_t) E_t \left[m_{t+1}^{\tilde{w}} m_{t+1}^{\tilde{\lambda}} p_{t+1}(n) \right]$$

$$m_{t+1}^{\tilde{w}} = \exp\left(-\tilde{w}_t' \varepsilon_{t+1} - \frac{\tilde{w}_t' \tilde{w}_t}{2} \right)$$

$$m_{t+1}^{\tilde{\lambda}} = \exp\left(-\tilde{\lambda}_t \varepsilon_{t+1} - \frac{\tilde{\lambda}_t' \tilde{\lambda}_t}{2} \right)$$

where E_t is taken with respect to the econometrician's model and $\tilde{w}_t = \tilde{w} x_t$ is the dubious investor's worst-case model.

Szőke's empirical strategy, II

- Assess improvements in predicted behavior of term structure of interest rates
- Use estimated worst-case dynamics to form distorted forecasts $\tilde{x}_{t+1} = (A - C\tilde{w})x_t$ and compare them to those of professional forecasters.
- Compute discounted relative entropy of worst-case twisted model $(A - C\tilde{w}), C, (D - G\bar{w}), G$ relative to the econometrician's model A, C, D, G and use it and Chernoff-Newman-Stuck entropy measures to assess difficulty of distinguishing two models.

Interpretations from Szőke model

Conditional mean distortions wx_t:

- PSS: w_t^* is vector of "mistakes" or "suboptimal forecasts"
- Szőke: $\tilde{w}_t$ is vector of "model uncertainty" prices: compensations that Szőke's representative agent charges to bear ε_{t+1} with unknown probability distribution

Insights from Szőke's model

- A theory of belief distortions $\tilde{w}_t = \tilde{w} x_t$
- A theory about the question that professional forecasters answer:
  - they answer with their worst-case model because they hear "tell me forecasts that rationalize your (max-min) decisions"
- A way to assess how large belief distortions are relative to the econometrician's model

Insights from Szőke's model, II

He uses his estimated Ξ matrix

- To infer an equivalence class of alternative parametric models parameterized by $\bar{w}$ that concerns the representative investor
  - It has more long-run risk than econometrician's model
- He infers a worst-case mean distortion $\tilde{w}x_t$ whose state dependence causes the term structure to move with x_t in ways that explain hitherto unexplained term structure movements (e.g., Shiller's "volatility puzzle")

Concluding remarks

Joint probability distributions of interest rates and macroeconomic shocks are important in macroeconomics

- Costs of aggregate fluctuations (business cycles)
- Consumption Euler equations (aka 'New Keynesian IS curves')
- Optimal taxation and government debt management
- Central bank 'expectations management' strategies
- Long-run risk (aka 'secular stagnation')

David Backus

Dave Backus contributed immensely and graciously to what we know and how we can go about learning more

Rather than curse darkness, Dave lit candles